MEASURING INDUSTRIAL STAGNATION: 
THE CASE OF THE US RAILROADS

ROGER L. M. DUNBAR AND MARSHALL SARNAT*

I. INTRODUCTION

Railroads in general, and those of the United States in particular, have been the subject of considerable controversy and debate for many years. At least four groups of economists have shown special interest in the economic and social role played by this often troubled industry.

Economic historians have debated for many years the extent of the railways' contribution to economic progress. Kmenta and Williamson [7], Neal [15] and Morgan [12] have tried to discover the factors which explain the investment behavior of the railroads. In addition, proponents of the 'new economic history' (for example, Fogel [3], Fogel and Engerman [4], and Fishlow [2]) have questioned the received wisdom which ascribes to the railroads a pivotal role in the process of economic growth in both Europe and America during the second half of the last century.¹ Specialists in industrial organization have evidenced a growing concern, especially in the United States, that the current levels of productivity and profitability in this key industry may be inadequate (Meyer and Morton [11]). Economists with an interest in regulation have long debated the extent (and desirability) of the regulatory authorities' impact on the competitive structure of US railroads (Morton [13]); and finally financial economists have commented on the low return to investors from railroad securities in the United States during the postwar period (Sarnat [20]).

At one time or another all of the above-mentioned classes of economists have expressed their fear that the poor performance of US railroads in recent years may have seriously impaired their ability to fulfill their proper economic function as public carriers of freight and passengers; the dramatic collapse of the Penn Central served to refocus attention on the structural problem of secular stagnation which has plagued the industry for many decades.

The purpose of the present paper is two-fold: first, we apply to the railroads the tools of modern risk analysis, using for this purpose a comparatively neglected data set—the long-term series of holding period rates of return on railroad shares—in order to derive an operational measure of secular stagnation. Clearly, US railroads constitute a propitious subject for such

* The authors gratefully acknowledge the assistance of Annemose Engelhardt, F. C. Heng and H. Lohman, as well as the helpful comments of a referee.

¹ For a critical review of the new approach to the history of transport, see O'Brien [17].
an experiment; few indeed would deny the existence of the serious structural imbalance which has plagued the railroads, culminating finally in the stagnation of this once prime growth industry. Second, the risk analysis will be used to gain additional insights into the industry's transition from growth to stagnation. The various stages of the industry's development are examined against the background of the data on rates of return.

Section II sets out the basic concepts of risk and return which underly the analysis, and then three alternative efficiency criteria are derived. The empirical results of the efficiency analysis are presented and interpreted in Section III; Section IV concludes the paper with a brief discussion of some of the implications of our findings.

II. BASIC DEFINITIONS AND COMPUTATIONAL PROCEDURES

A. Rate of Return

Data on share prices and dividend yields are available for three categories of US common stock, industrials, utilities and railroads, for a period of more than 100 years. As a first step in the present analysis, annual rates of return were calculated for each of these industries, using the relevant index of common stock prices for each industry.

The annual rate of return on the common stock of a given industrial group was defined as follows:

\[
 r_t = \frac{D_t + (P_t - P_{t-1})}{P_{t-1}} = \frac{D_t + P_t}{P_{t-1}} - 1
\]

(1)

where

- \( D_t \) = the annual dividend;
- \( P_t \) = the relevant index of share prices for year \( t \);
- \( r_t \) = the rate of return in year \( t \).

In the actual calculations, dividend yields were used to approximate current dividend payments:

\[
 r_t = \frac{P_t}{P_{t-1}} (1 + T_t)
\]

(2)

where \( T_t \) denotes the current dividend yield on the shares which comprise the relevant share price index of the industry in question.

---

2 The data on share prices and dividend yields were obtained from Historical Statistics of the United States: Colonial Times to 1970, the Federal Reserve Bulletin and Moody's Investors Service. In all cases the share price indexes and dividend yields refer to annual averages.

3 All calculations are based on annual averages of the share price index.

4 The use of the index of share prices, rather than an average of the stock prices themselves, helps to overcome a problem which has hampered the empirical study of rates of return. Cash dividends and the change of stock prices do not necessarily, and usually do not, represent the total return to shareholders, owing to the existence of a wide variety of stock dividends, splits and pre-emptive rights issues. Fortunately, for our purposes, the published share price index is adjusted for such capital alterations.
The mean (expected) rate of return, \( \bar{r} \) for each category of shares, was calculated by taking the arithmetic average of the annual returns:

\[
\bar{r} = \frac{1}{N} \sum_{t=1}^{N} r_t
\]

Similarly, the compounded average annual rate of return, \( \bar{r}_g \), was calculated for each industry using the familiar formula for the geometric mean:

\[
\bar{r}_g = N \sqrt[N]{\prod_{t=1}^{N} (1 + r_t)} - 1
\]

This formula implies the inter-year reinvestment of the annual dividends.\(^5\)

Since inflation is an important variable affecting the rate of return, nominal returns have been adjusted (where necessary) to reflect the impact of changes in the general level of prices. The inflation-adjusted rate of return, i.e., in terms of real purchasing power, was found by deflating the nominal annual rates of return, using for this purpose the consumer price index:

\[
R_t = \frac{1 + r_t}{1 + g_t}
\]

where

- \( R_t \) = the real rate of return in year \( t \).
- \( g_t \) = the percentage change in the consumer price index during the year \( t \).

B. Risk

The efficiency analysis of choice involving risk identifies the latter with the variability of returns, i.e., with the dispersion (both positive and negative) of returns from the expected (mean) return. The variability of the rates of return was calculated by taking the standard deviation of returns:

\[
\sigma = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (r_t - \bar{r})^2}
\]

where \( \bar{r} \) and \( \sigma \) denote the arithmetic average and standard deviation of the annual returns respectively.

C. Efficiency Criteria

In their monumental work on the theory of games, von Neumann and Morgenstern [23] provided a rigorous axiomatic justification for the use of expected utility to explain rational choice under conditions of uncertainty, by demonstrating that if a number of apparently reasonable consistency requirements are satisfied, utility can be introduced in a way that ensures that rational choice among risky alternatives will be made solely on the basis of their expected utilities.

\(^5\) The reader should note, however, that the formula ignores the intra-year reinvestment of cash dividends, i.e., the timing of dividend receipts during a given year.
The von Neumann–Morgenstern analysis led directly to the formulation of efficiency criteria based on expected utility, such a criterion being defined as a decision rule for partitioning all potential investment options into two mutually exclusive sets: an efficient set of potentially acceptable options, and an inefficient set. No expected utility maximizer will willingly choose an option out of the inefficient set.

The modern theory of investment choice uses such criteria to dichotomize the decision process into two steps: first, the number of options is reduced by constructing the efficient set of alternatives using an efficiency criterion which is appropriate for a given class of utility functions (for example, the class of all risk aversers). Second, the individual is assumed to make his final choice out of the reduced set in accordance with the shape of his specific utility function. Now, let us examine the properties of three popular efficiency criteria.

1. The Expected Mean-variance (E–V) Criterion

The efficiency analysis of choice under conditions of uncertainty which emerged following the seminal works of Harry Markowitz [9], [10] and James Tobin [22] has demonstrated that in a number of significant cases risk-averse individuals who maximize their expected utility invariably choose an investment option which is efficient in terms of the first two moments of the distribution of returns. The popular Markowitz E–V criterion for screening inefficient (dominated) options states that an option A will dominate an option B, thereby eliminating B from the efficient set, if the following inequalities hold:6

\[ E(x_a) \geq E(x_b) \quad \text{and} \quad \text{Var}(x_a) \leq \text{Var}(x_b). \]

Tobin has shown that this criterion provides an appropriate screening device if we assume quadratic utility, or alternatively, if we assume concave utility (i.e., risk aversion) and normally distributed distributions of returns.7

The major advantage of the E–V criterion lies in its simplicity: it permits us to focus attention solely on the first 2 moments of the distribution of outcomes (expected mean and the variance).

However, two serious objections have been raised regarding its use. First, the assumption of quadratic utility raises difficulties. Such utility functions are only valid for a bounded range; Arrow [1] and Pratt [18] have shown that the unrestricted use of quadratic utility implies the economically unacceptable assumption of ever-increasing absolute risk aversion. Second, for many problems including the one at hand, the alternative assumption

---

6 On the additional condition that at least one of the strong inequalities holds.
7 Levy and Sarnat [8] have shown that the E–V criterion also provides an appropriate decision-rule in the more general case of all independent two-parameter distributions, but on the further condition that non-intersecting cumulative frequency distributions are eliminated.
of normal, or near normal, frequency distributions of returns often does not hold.

2. **First Degree Stochastic Dominance (FSD)**

These shortcomings of the $E-V$ approach have led to the development of more general ‘distribution-free’ efficiency criteria, two of which will be utilized in the analysis below. In the most general case, in which no restrictions are placed on investors’ utility functions beyond the reasonable assumption that they be non-decreasing with respect to returns, $x$, (i.e., the analysis is appropriate for both ‘risk lovers’ and ‘risk averters’), an optimal efficiency criterion can be defined as follows: given two cumulative probability distributions $F$ and $G$, option $F$ will dominate a second option $G$ (thereby eliminating the latter from the efficient set), independent of the concavity or convexity of the utility function, *if and only if*:

$F(x) \leq G(x)$

for all values of $x$, with a strong inequality $F(x_0) < G(x_0)$ holding for at least one value of $x$.\(^8\) This is tantamount to assuming that the two cumulative probability distributions do not intersect, or in other words, option $F$ stochastically dominates option $G$.

This criterion can be illustrated by a simple graphical device. Figure 1 plots the cumulative probability functions of two investment options, $A$ and $B$. As $B$ lies to the right of $A$, the FSD criterion is satisfied and option $A$ is eliminated from the efficient set, independent of the shape of individuals’ utility functions.

Upon reflection, the stochastic dominance analysis is almost intuitively obvious. The FSD criterion, which has been defined in terms of cumulative probability distributions, stipulates that any option $F$ dominates another option $G$ if $F(x) \leq G(x)$. But this is equivalent to the requirement that the probability of achieving a return lower than some amount, say $k$, will always be smaller for option $F$ than for option $G$. Since $F(k) \leq G(k)$ by definition, it follows that

$Pr_F (x \leq k) \leq Pr_G (x \leq k)$

where $Pr$ denotes probability.\(^9\)

3. **Second Degree Stochastic Dominance (SSD)**

The FSD criterion is very general; it applies very weak restrictions on

---

\(^8\) For a formal proof see Hadar and Russell [5] and Hanoch and Levy [6].

\(^9\) Multiplying both sides of (7) by $-1$ and adding unity, we obtain $1-F(x) \geq 1-G(x)$. This expression is equivalent to the probability of receiving a return which is greater than or equal to a given return, $x$. Since the FSD criterion must hold for all levels of return, this implies that the probability of receiving a return greater or equal to some level, $k$, must always be higher for option $F$ than for option $G$. 

utility functions, which are only required to have non-negative first derivatives, and, therefore, the criterion is appropriate for all individuals independent of the pattern of their risk-return tradeoffs. In almost all economic analysis, however, universal risk-aversion is assumed to prevail, i.e., individuals' utility functions are assumed to be concave throughout the relevant range.

Given this realistic assumption, an efficiency criterion for the class of all risk averters, i.e., Second Degree Stochastic Dominance (SSD) criterion, can be defined. Again denoting the cumulative probability distributions of two options by $F$ and $G$, a necessary and sufficient condition for $F$ to dominate $G$ for all risk averters (concave utility functions) is that

$$\int_{-\infty}^{\infty} [G(t) - F(t)] \, dt \geq 0$$

10 For empirical evidence on investors' behavior in support of this statement see Levy and Sarnat [8, ch. 9].
with a strong inequality holding for at least one value of $x$. Thus, unlike FSD, SSD permits the cumulative probability distributions to intersect, on the condition that the cumulative difference between $G$ and $F$ remains non-negative over the entire domain of $x$.

Once again, the efficiency criterion can be illustrated graphically. Figure 2 plots the cumulative probability distributions of two alternatives, $F$ and $G$. Although the two functions intersect several times, it is clear by inspection that the cumulative first differences between the two functions remain positive for all values of return ($x$), that is, the cumulative shaded areas for which $G(x) > F(x)$ always exceeded the cumulative unshaded areas for which $G(x) < F(x)$ over the entire domain of $x$. Hence, option $F$ dominates $G$.

\[\text{Figure 2. Second Degree Stochastic Dominance: A Graphical Illustration}\]

---

11 This means that given the stipulation that the first two derivatives of the utility function be non-negative, and not positive, respectively (i.e., $U'(x) \geq 0$ and $U''(x) \leq 0$), no criterion can be found to further reduce the SSD efficient set without placing additional restrictions on investors’ utility functions and/or on the probability distribution of returns. For formal proof of the optimality of the SSD criterion, see Hadar and Russell [5] and Hanoch and Levy [6].
according to the SSD criterion, and $G$ is eliminated from the efficient set for all risk averters.

Before going on to the empirical analysis, three things should be emphasized:

(a) The two stochastic dominance criteria, FSD and SSD, permit an analysis of the risk-return relationship without recourse to restrictions on the frequency distributions of monetary returns. Thus, unlike the $E-V$ criteria, the distributions of returns need not be normal, or for that matter even symmetrical.

(b) The FSD and SSD criteria are essentially complements rather than substitutes. In most cases, it is desirable to apply both the stronger, SSD, and the weaker, FSD, criteria.

(c) The SSD efficient set is a subset of the FSD efficient set.

The stronger assumptions of the former permit a more sensitive screening of options; hence the empirically derived SSD efficient set can be expected to be smaller, and in many cases substantially smaller, than its FSD counterpart.

III. EMPIRICAL RESULTS

Now let us apply the efficiency criteria to the analysis of the risk-return relationship of the US railroads. The construction of a continuous and comparable series of rates of return on the investment in equities over a sufficiently long period of time was simplified by recourse to the device of using share price indices rather than individual share prices. This approach permits the calculation of annual (average) rates of return for three categories of common stock: industrials, railroads and utilities over the period 1872–1974.\footnote{In the analysis which follows the rates of return for the period 1872–1929 have been calculated on the basis of the share price indices and divided yields which appear in \textit{Historical Statistics of the United States: Colonial Times to 1970}; for subsequent years the data were obtained from \textit{Moody’s Industrial Manual}.}

Nominal and real rates of return and their standard deviations for selected cumulative periods from 1872–1974 are given in Table I. Even a cursory examination of the data suffices to show that the investment in industrial shares has earned a significantly higher rate of return over the entire period. The relationship between railroad and utility returns is somewhat more complex. Over the entire period, 1872–1974, the real annual average return of railroad shares slightly exceeded that of the utilities, 4.11\% as compared with 3.92\%. However, this relationship does not hold for the earlier period; for example, the long-term rate of return on utilities was higher than that of railroads in each of the subperiods 1872–1900, 1872–1929 and 1872–1946.

Although the mean return on railroad investment has been positive, clearly this does not necessarily imply adequate economic performance or
### Table I

<table>
<thead>
<tr>
<th>Industrial branch</th>
<th>Nominal rate of return</th>
<th>Standard deviation</th>
<th>Real rate of return</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872–1974</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrials</td>
<td>8.85</td>
<td>17.54</td>
<td>7.92</td>
<td>17.74</td>
</tr>
<tr>
<td>Railroads</td>
<td>5.90</td>
<td>18.39</td>
<td>4.11</td>
<td>18.74</td>
</tr>
<tr>
<td>Utilities</td>
<td>5.79</td>
<td>17.53</td>
<td>3.92</td>
<td>18.30</td>
</tr>
<tr>
<td>1872–1946</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrials</td>
<td>8.24</td>
<td>18.68</td>
<td>7.06</td>
<td>18.56</td>
</tr>
<tr>
<td>Railroads</td>
<td>5.20</td>
<td>18.96</td>
<td>4.05</td>
<td>19.19</td>
</tr>
<tr>
<td>Utilities</td>
<td>5.50</td>
<td>19.22</td>
<td>4.34</td>
<td>19.70</td>
</tr>
<tr>
<td>1872–1929</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrials</td>
<td>9.84</td>
<td>15.70</td>
<td>8.53</td>
<td>16.68</td>
</tr>
<tr>
<td>Railroads</td>
<td>7.00</td>
<td>12.93</td>
<td>5.73</td>
<td>14.10</td>
</tr>
<tr>
<td>Utilities</td>
<td>7.93</td>
<td>14.81</td>
<td>6.65</td>
<td>16.71</td>
</tr>
<tr>
<td>1872–1900</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrials</td>
<td>7.53</td>
<td>12.24</td>
<td>8.46</td>
<td>11.64</td>
</tr>
<tr>
<td>Railroads</td>
<td>5.64</td>
<td>12.76</td>
<td>6.56</td>
<td>12.27</td>
</tr>
<tr>
<td>Utilities</td>
<td>6.94</td>
<td>12.33</td>
<td>7.87</td>
<td>12.26</td>
</tr>
</tbody>
</table>

*a Geometric mean


the absence of stagnation. Concentration on a single parameter ignores the crucial role played by risk.\(^{13}\)

Since an examination of the distributions of return indicates a rather pronounced degree of asymmetry, we shall incorporate risk into the analysis by applying the FSD and SSD efficiency criteria to the empirically generated distributions of nominal and real rates of return. The results of this analysis are presented in Table II. Railroad shares constitute an efficient alternative, in the FSD sense, both in nominal and real terms, for all of the cumulative empirical distributions of returns, starting with that for 1872–90 and ending with the distribution for the period 1872–1974.\(^{14}\) This is not surprising, considering the weakness of the FSD criterion’s assumptions.

This is not the case with respect to the SSD, i.e., the appropriate criterion for the class of all risk averters. If, for convenience, we focus our attention

---

\(^{13}\) Moreover, the single parameter analysis is ‘data dependent’. When the Standard and Poor’s share price index, (which is the index appearing in Historical Statistics) is substituted for the Moody’s index after 1929, the relationship between the rates of return on railroad and utility shares is reversed for the overall period 1872–1974. The relevant figures for the annual average real rate of return are 4.70% for the utilities and 4.33% for the railroads.

\(^{14}\) Since data on the rates of return are available, starting in 1872, and the distributions can only be estimated from a reasonably large number of years, the period 1872–90 is the earliest period for which the analysis was attempted.
Table II
RELATIVE EFFICIENCY OF RAILROAD SHARES,
ESTIMATED FROM NOMINAL AND REAL RETURNS
SELECTED PERIODS, 1872–1974

<table>
<thead>
<tr>
<th>Time period</th>
<th>FSD criterion</th>
<th>SSD criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872–1890</td>
<td>Efficient</td>
<td>Not Efficient</td>
</tr>
<tr>
<td>1872–1902</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872–1903</td>
<td>Efficient</td>
<td>Efficient</td>
</tr>
<tr>
<td>1872–1937</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872–1938</td>
<td>Efficient</td>
<td>Not Efficient</td>
</tr>
<tr>
<td>1872–1974</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. REAL RETURNS

<table>
<thead>
<tr>
<th>Time period</th>
<th>FSD criterion</th>
<th>SSD criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872–1890</td>
<td>Efficient</td>
<td>Not Efficient</td>
</tr>
<tr>
<td>1872–1901</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872–1902</td>
<td>Efficient</td>
<td>Efficient</td>
</tr>
<tr>
<td>1872–1934</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872–1935</td>
<td>Efficient</td>
<td>Not Efficient</td>
</tr>
<tr>
<td>1872–1974</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

upon the real analysis, a stable pattern emerges. Using ex post data on the rates of return, the railroads were not efficient (i.e., they were dominated in the sense of SSD) during the period 1872–1901, were efficient during the years 1902–34, and were once again inefficient from 1935–74.

These results confirm the independent analysis of the railroads' difficulties. Although the 'SSD inefficiency' of the railways in the latter part of the 19th century coincides with a period of expansion, it was also marked by a spate of financial crises and bankruptcies which peaked in 1873. Although the railroads continued to expand, the Panic of 1893 led to a substantial reorganization of the railroads' financial structures, and at one time, 25% of
the nation's total railroad mileage was in receivership. (Neal [15]). This early period of stagnation also coincides with the rise of a broad national market for industrial securities (Navin and Sears [14] and Nelson [16]) which led some investors to search elsewhere and eventually turn to industrial trials for more satisfactory returns.

It was not until the beginning of the 1930s that the long-term stagnant nature of the industry became fully apparent. At the end of the 1920s and beginning of the 1930s, the growth rate of railroad assets fell drastically (see Figure 3); in addition, the flow of new patents and inventions also declined after the 1930s and has not recovered (Schmookler [21]). Thus, the prolonged inefficient performance of railroad shares, as measured by the SSD criterion, correctly mirrors the secular stagnation of this vital industry.

IV. SOME IMPLICATIONS OF THE ANALYSIS

Subject to the usual caveat regarding the use of ex post data as proxies for ex ante expectations, our findings suggest that the stochastic dominance analytical technique can be used to identify and analyze cases of secular stagnation. SSD inefficiency which persists for a long period of time implies the kind of fundamental, rather than transitory, disequilibrium which is characteristic of secular stagnation. The stochastic dominance criterion has the further advantage of being a 'distribution free' measure of inefficiency. Thus it is an appropriate criterion for the negatively skewed distributions which are characteristic of stagnating firms or industries.

The analysis also suggests the persistance of market disequilibrium, in the sense that the fall in railroad share prices has not been sufficient to restore such shares (taken as a group) to the set of efficient alternatives. But this finding also raises some serious questions. If the market performance of the railroad industry in terms of risk and return has been so inferior over such a long time frame, how does the industry survive? Or—more specifically—how in the face of such performance did the industry succeed in raising its not inconsiderable capital requirements? One possible answer lies in the direction of the regulated nature of the US railroads. Government legislation,

---

15 The fierce competition and continuing heavy capital needs for expansion which were characteristic of this period resulted in a new wave of bankruptcies during the 1880s and early 1890s, particularly of many large roads which were still in the process of construction. One of the results of these bankruptcies was the establishment of the Interstate Commerce Commission through the Act to Regulate Commerce in 1887.

16 The turning point can be refined somewhat by examining the measure of the distributions' asymmetry. In the case of the railroads, all of the distributions starting with 1872–90 have positive asymmetry until that of 1872–1930. All of the subsequent distributions have negative asymmetry.

17 Using analytical techniques Sahal [15] has demonstrated that this decline in railroad innovations constitutes a clear break with the past.

18 Portfolio considerations offer little help in this context. The poor performance of railroad shares was not offset by their relatively low covariance with the other industry groups; see Sarnat [20].
and the provisions for railroad reorganizations in particular, have undoubtedly played a major role in preventing the industry’s collapse. However, it would appear that the US railroads constitute a special case in which financial innovation has served as a sort of survival mechanism. Plagued as they have been throughout much of their history by bankruptcies and financial crises, the railroads very early found it difficult to finance their capital requirements by conventional means. In order to circumvent their problem, the equipment trust certificate, an investment instrument which effectively shields the investor from the riskiness of enterprise, was introduced. Thus, despite the inadequate return on capital investment, the railroads have been able to raise debt capital at relatively low-risk premiums.

NEW YORK UNIVERSITY AND THE
HEBREW UNIVERSITY, JERUSALEM

REFERENCES


